

# NISRA CORONAVIRUS (COVID-19) OPINION SURVEY

## TECHNICAL REPORT

### Background

On 20 April 2020, NISRA launched a new Coronavirus (COVID-19) Opinion Survey designed to measure how the Coronavirus (COVID-19) pandemic was affecting peoples' lives and behaviour in Northern Ireland.

The NISRA Coronavirus (COVID-19) Opinion Survey questionnaire was based on a similar survey that was being conducted by the Office for National Statistics (ONS) in Great Britain.

<https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandwellbeing/bulletins/coronavirusandthesocialimpactsongreatbritain/previousReleases>

ONS provided NISRA with each of the questionnaires they had used in their weekly surveys. The questions asked in the ONS weekly surveys evolved throughout the pandemic period, however, there was a degree of consistency with the majority of questions being retained in each version of the questionnaire.

NISRA carries out its Coronavirus (COVID-19) Opinion Survey as a weekly survey with the questionnaire being refreshed every four weeks. Each four week period is, therefore, considered to be a new phase of the survey. The NISRA COVID-19 questionnaire used in Phase 1 was based on the OPN3 and OPN4 questionnaires used by ONS and Phase 2 was based on the OPN7 questionnaire.

NISRA is currently planning to carry out five phases of the survey.

## Sampling Method

The weekly samples for the NISRA Coronavirus (COVID-19) Survey consisted of a systematic random sample of 1,000 addresses selected from the NISRA Address Register (NAR). The NAR is developed within NISRA and is primarily based on the Land & Property Services (LPS) database.

Pointer is the address database for Northern Ireland maintained by Land & Property Services (LPS). People living in institutions (though not in private households in such institutions) were excluded.

## Data Collection mode

The interviews were carried out using CATI (computer assisted telephone interview). NISRA does not hold any details relating to persons living at the sampled addresses or telephone contact details. Therefore, NISRA wrote out to the sampled addresses and asked them to contact the Agency by telephone or online to organise/book a convenient time for the interview to take place. The interviews were conducted by NISRA Survey Interviewers.

## Respondent Selection

While the NAR provides a good sampling frame of addresses, it contains no information about the number of people living at an address. Further selection stages are therefore required to convert the listing of addresses to a listing of individuals from which one person (the 'selected respondent') is chosen to complete the questionnaire.

The interviewers list all members of the household who are eligible for inclusion in the sample: that is, all persons currently aged 16 or over living in the household. From this listing of eligible adults, the computer selects one adult by a random selection procedure. This person, the selected respondent, is then asked to complete the interview. Should the selected respondent refuse to participate or if no contact can be made with them, the interviewer can accept no other replacement for that person.

## Fieldwork

The results in this release relate to Phase 1 and Phase 2 of the survey. These are cases sampled in the first eight weeks of the survey. Further results relating to subsequent phases will be published periodically when the data becomes available.

In total, 1,785 persons from Phase 1 and Phase 2 participated in the survey. The interviews took place in the period 21 April to 23 July 2020. This represents a survey response rate of 24% of eligible addresses. A breakdown of responses to the survey is provided in Table 1.

**Table 1 Survey Response Phase 1 & Phase 2 (Weeks 1-8)**

<i>All respondents</i>	<i>Number</i>	<i>Percentage (%)</i>
<i>Number of addresses sampled</i>	<b>8000</b>	
<i>Eligible (after adjustment)<sup>1</sup></i>	7360	100
<i>Did not respond to advance letter</i>	5339	73
<i>Number of households responding to invitation</i>	2018	27
<i>Number of persons interviewed</i>	1785	
<i>Partially co-operating</i>	0	0
<i>Total co-operating</i>	<b>1785</b>	24
<i>Refusal to co-operate by selected respondent<sup>2</sup></i>	193	3
<i>No contact with selected respondent</i>	40	1

<sup>1</sup> The adjusted eligible households assumes that 8% of addresses would fall into the category of ineligible for the survey because no-one lives at the address. This outcome mainly covers vacant properties, second homes, demolished or derelict properties or properties under construction. Ineligible addresses are normally validated by interviewers in the field when they visit an address. The absence of interviewers in the field means that we cannot verify ineligible addresses for the NISRA COVID-19 survey. However, we have assumed an 8% ineligible rate in the COVID-19 weekly samples as this was the typical rate of ineligibles experienced in 2019-20 fieldwork across multiple surveys.

<sup>2</sup> In multi-person households, the person selected for the survey from the household grid may not be the same person who expressed an interest in participating in the survey in the first place.

## Weighting

Selecting only one individual for interview at each sampled address means that the probability of selection for the survey is inversely related to the size of the household. In other words individuals living in large households have a lower chance of being included in the sample than individuals in small households.

Before analysis, the data are weighted in relation to the number of eligible adults at the address, derived from the details of the household structure recorded by interviewers on the questionnaire. This is known as a design weight and this weighting process adjusts the results

to those which would have been achieved if the sample had been drawn as a systematic random sample of adults rather than of addresses.

Further weighting was applied to the data to adjust for differences in the composition of the achieved sample (after the design weight was applied) compared to the same characteristics in the population, known as non-response weights. The non-response rates were applied sequentially to calibrate the characteristics of the achieved sample to the population distribution in terms of sex by gender and then by Local Government District (LGD14).

Once the data was weighted by both the design weight and the non-response weights, the achieved sample became more representative of the population the survey aimed to reflect. Annex 1 provides a breakdown for a number of demographic characteristics which show the impact of non-weighting and weighting on the demographic composition of the achieved sample.

All figures published in the Phase 1 and Phase 2 Results are weighted estimates.

## Sampling Error

No sample is likely to reflect precisely the characteristics of the population it is drawn from because of both sampling and non-sampling errors. An estimate of the amount of error due to the sampling process can be calculated. For a simple random sample design, in which every member of the sampled population has an equal and independent chance of inclusion in the sample, the sampling error of any percentage,  $p$ , can be calculated by the formula:

$$s.e.(p) = \sqrt{(p * (100 - p)/n)}$$

Where  $n$  is the number of respondents on which the percentage is based.

The sample for the NISRA COVID-19 Survey is drawn as a systematic random sample, and this formula can be used to calculate the sampling error of any percentage estimate from the survey.

A confidence interval for the population percentage can be calculated by the formula 95 per cent confidence interval =  $p \pm 1.96 * s.e.(p)$

If 100 similar, independent samples were chosen from the same population, 95 of them would be expected to yield an estimate for the percentage,  $p$ , within this confidence interval.

Selecting one individual for interview at the sampled address means that each person in the sample has a different probability of selection, which depends on the size of household they live in. A design effect therefore exists in the survey data. The design effect for the NISRA COVID-19 Opinion Survey Phase 1 & Phase 2 results (based on 1785 responses) was calculated to be 1.2 and the design factor is 1.1. The design factor has been used to adjust the standard error and confidence intervals of the Phase 1 & Phase 2 survey estimates.

Table 2 provides confidence intervals for a range of percentages based on the number of achieved cases in Phase 1 and Phase 2 ( $n=1785$ ).

For example, if we assume a survey estimate for a particular variable in the NISRA Coronavirus (COVID-19) Opinion survey is 20%. Based on the number of people participating in Phases 1 & 2 ( $n=1785$ ), we can be 95% confident that the true population measure is  $\pm 2.0\%$  of that survey estimate. In other words, we can be 95% confident that the true population figure lies in the range 18.0%-22.0%.

**Table 2 – Illustrative confidence intervals for Phase 1 & Phase 2 survey estimates**

<b>Range of error (+/-) for 95% CI - for range of percentages by achieved sample size on NISRA Coronavirus (COVID-19) Opinion Survey</b> (assumes a systematic random sample with the design factor applied)					
<b>Range</b>	<b>5% or 95%</b>	<b>10% or 90%</b>	<b>20% or 80%</b>	<b>30% or 70%</b>	<b>50%</b>
<i>Base (1785) cases</i>	1.1	1.5	2.0	2.3	2.6

## Significant Differences between estimates

Any differences referred to in Phase 1 & Phase 2 results as “statistically significant” are based on non-overlapping confidence intervals using the 95% confidence intervals. Confidence intervals and comparisons of confidence intervals have been made on unrounded data.

## Strengths and Limitations

### **The main strengths of the NISRA Coronavirus (COVID-19) Opinion Survey include:**

It meets data needs. The questionnaire was developed by ONS and included customer consultation. Furthermore, questionnaire design expertise was applied in the development stages. In Northern Ireland, further consultation on the Phase 1 & Phase 2 questionnaires took place with NI Departments.

Robust methods were adopted for the survey's sampling and weighting strategies to limit the impact of bias.

The data was collected with a Blaise 5 CATI questionnaire which has set ranges for closed questions, automated routing and internal consistency checks with reduces interviewer error and data input error.

Quality assurance procedures are undertaken throughout the analysis stages to minimise the risk of error.

The data was collected by NISRA Survey Interviewers who are highly skilled and trained and who have a long track record of collecting high quality survey data.

**The main limitations of the NISRA Coronavirus (COVID-19) Opinion Survey include:**

The achieved sample size of 1785 responses is quite small and limits the amount of sub-analysis which can be carried out on the data.

The number of responses to the survey (n=1785) affects the precision of survey estimates. The level of sampling error associated with the NISRA Coronavirus (COVID-19) Opinion Survey estimates may be as high as +/- 2.6% for some variables.

In any survey there is the possibility of non-response bias. Non-response bias arises if the characteristics of non-respondents differ significantly from those of respondents in such a way that they are reflected in the responses given in the survey. Accurate estimates of non-response bias can only be obtained by comparing characteristics of the achieved sample with the distribution of the same characteristics in the population at the time of sampling. Such comparisons are usually made to Census of Population Estimates.

Weighting for non-response (post-stratification) as applied to the Phase 1 & Phase 2 NISRA Coronavirus (COVID-19) Opinion Survey results should reduce bias, but it must be acknowledged that it will not eliminate bias. The reasons individuals choose to take part in surveys are complex and depend on lots of factors specific to the individual. As a result, the non-response biases in surveys are likely to be complex. Post-stratification works on the assumption that, by aligning the survey to the population along a small number of dimensions, such as age and gender or local government district, many of these complex biases will reduce. However, it would be misleading to suggest that they will be eliminated.

## Annex 1

**Table A1 Characteristics of the achieved sample by age and sex**

<i>All respondents</i>	<i>Non-Weighted</i>	<i>Weighted*</i>	<i>2019 MYE</i>
16-24	6.6	13.3	13.7
25-34	10.3	16.6	16.6
35-44	13.8	16.5	16.1
45-54	18.4	17.3	17.2
55-64	19.3	15.3	15.4
65+	31.7	21.0	21.0
<b>Total</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>
<b>Unweighted base</b>	<b>1785</b>	<b>1785</b>	<b>1497742</b>
<b>Males</b>			
16-24	3.0	7.0	7.1
25-34	4.1	8.1	8.3
35-44	6.1	8.0	7.8
45-54	7.7	8.6	8.4
55-64	8.2	7.3	7.6
65+	14.6	9.5	9.6
<b>Total</b>	<b>43.8</b>	<b>48.4</b>	<b>48.7</b>
<b>Unweighted base</b>	<b>782</b>	<b>782</b>	<b>729636</b>
<b>Females</b>			
16-24	3.6	6.4	6.6
25-34	6.1	8.5	8.3
35-44	7.7	8.5	8.2
45-54	10.6	8.7	8.8
55-64	11.0	8.0	7.8
65+	17.1	11.5	11.5
<b>Total</b>	<b>56.2</b>	<b>51.6</b>	<b>51.3</b>
<b>Unweighted base</b>	<b>1003</b>	<b>1003</b>	<b>768106</b>

\*weighted for survey design, age by sex and then LGD2014

**Table A2 Characteristics of the achieved sample by LGD2014**

	<i>Non-Weighted</i>	<i>Weighted*</i>	<i>2019 MYE 16+ (Person Level)</i>
ANTRIM & NEWTOWNABBEY	8.7	7.6	7.6%
ARMAGH CITY, BANBRIDGE & CRAIGAVON	11.3	11.2	11.2%
BELFAST CITY	17.7	18.4	18.4%
CAUSEWAY COAST & GLENS	8.0	7.8	7.8%
DERRY CITY & STRABANE	6.8	7.9	7.9%
FERMANAGH & OMAGH	4.8	6.1	6.1%
LISBURN & CASTLEREAGH CITY	9.9	7.8	7.8%
MID & EAST ANTRIM	8.2	7.5	7.5%
MID-ULSTER	4.9	7.6	7.6%
NEWRY, MOURNE & DOWN	8.9	9.4	9.4%
ARDS & NORTH DOWN	10.9	8.8	8.8%
<b>Unweighted base</b>	<b>1785</b>	<b>1785</b>	<b>1497742</b>

\*weighted for survey design, age by sex and then LGD2014

**Table A3 Characteristics of the achieved sample by ILO Status**

	<i>Non-Weighted</i>	<i>Weighted*</i>	<i>LFS (Feb-Apr 2020)</i>
Employed	49.9	58.6	58.9
Unemployed	1.1	1.4	2.3
Inactive	49.0	40.0	39.8
<b>Unweighted base</b>	<b>1785</b>	<b>1785</b>	<b>4908</b>

\*weighted for survey design, age by sex and then LGD2014

**Table A4 Characteristics of the achieved sample by Religious Denomination**

	<i>Non-Weighted</i>	<i>Weighted*</i>	<i>CHS 2018/19</i>
Catholic	35.0	39.3	42.7
Protestant	49.4	44.2	44.9
Other or no religion	15.0	16.2	11.3
Unwilling to answer/can't be determined	0.6	0.4	1.1
<b>Unweighted base</b>	<b>1785</b>	<b>1785</b>	<b>10902</b>

\*weighted for survey design, age by sex and then LGD2014